

PCT/US 99/04376

IPEA/US 22 MAY 2000

PCT/US99/04376

49  
75  
Claims1  
2  
3  
4 What is claimed:5 1. An invention substantially as described in the description.  
6 2. An invention substantially as described and illustrated in the description.  
7 3. A process for identifying one or more bi-allelic markers linked to a bi-allelic genetic characteristic  
8 gene in a species of creatures, comprising the steps of:9  
10 a) choosing two or more bi-allelic covering markers so that a CL-F region is systematically covered by  
11 the two or more covering markers, the CL-F region being a collection of points on a two-dimensional  
12 plane, the two-dimensional plane having the two orthogonal dimensions of chromosomal location and  
13 least common allele frequency;14  
15 b) choosing a statistical linkage test based on allelic association for each covering marker;16  
17 c) choosing a sample of individuals for each covering marker;18  
19 d) obtaining genotype data/sample allele frequency data for each covering marker and the sample  
20 chosen for each covering marker, and obtaining phenotype status data for the genetic characteristic for  
21 each individual in the sample chosen for each covering marker;22  
23 e) calculating evidence for linkage between each covering marker and the gene using the statistical  
24 linkage test based on allelic association chosen for each covering marker and the genotype  
25 data/sample allele frequency data for each covering marker and using the phenotype status data for the  
26 genetic characteristic for each individual in the sample chosen for each covering marker obtained in d);  
27 and28  
29 f) identifying those covering markers as linked to the genetic characteristic gene which show evidence  
30 for linkage based on the calculations of step e).31 2. A process as in claim 1 wherein the CL-F region is N covered to within a CL-F distance  $\delta$  by the two  
32 or more bi-allelic covering markers, so that each point in the region is within the CL-F distance  $\delta$  of N or  
33 more of the covering markers, wherein  $\delta$  is equal to about  $[\delta_{CL}, 0.25]$  or the equivalent thereof,  $\delta_{CL}$  is  
34 equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium hasAMENDED SHEET  
50

50  
76

1 been observed between any polymorphisms in any population of the species, N is an integer greater  
2 than or equal to 1.

3 <sup>2</sup> 3. A process as in claim 1, wherein the CL-F region includes 81 percent or more of the centerpoints of  
4 the matrix centerpoint lattice of a CL-F matrix, the number of cells in the matrix being greater than or  
5 equal to three, wherein the matrix has R rows and C columns, each cell of the matrix being of length  $L_{MC}$   
6 and width  $W_{MC}$ , and  $L_{MC}$  being less than or equal to about  $2\delta_{CL}$ , and  $W_{MC}$  being less than or equal to 0.5,  
7  $\delta_{CL}$  is equal to the largest chromosomal length, computed by any method, for which linkage  
8 disequilibrium has been observed between any polymorphisms in any population of the species, there  
9 being N or more covering markers in each cell of 81 percent or more of the cells of the matrix, N is an  
10 integer greater than or equal to 1; the covering markers being distributed over a chromosomal region of  
11 interest, the region of interest being approximately the smallest chromosome interval that contains all of  
12 the covering markers, and the covering markers comprising essentially less than all of the  
13 polymorphisms in the region of interest.

14 6. A process as in claim 5, wherein the covering markers are substantially nonevenly distributed across  
15 a chromosome or a chromosomal segment.

16 7. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a  
17 chromosome or a chromosomal segment, and wherein there is a subgroup of one or more of the  
18 covering markers, and each of the markers in the subgroup is chosen without substantial preference for  
19 the least common allele frequency of each of the markers in the subgroup being close to 0.5, and the  
20 number of covering markers in the subgroup is about 5 percent or more of the total number of covering  
21 markers.

22 <sup>3</sup> 4. A process as in claim 3, wherein the covering markers are substantially evenly distributed across a  
23 chromosome or a chromosomal segment, and wherein there is a subgroup of one or more of the  
24 covering markers, and each of the markers in the subgroup is chosen without substantial preference for  
25 the least common allele frequency of each of the markers in the subgroup being close to 0.5.

26 9. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a  
27 chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance of  
28 the covering markers is greater than 2 cM or the equivalent thereof and the least common allele  
29 frequency of one or more of the covering markers is less than 0.3, or wherein (2) the least common  
30 allele frequency of one or more of the covering markers is less than 0.2, or wherein (3) the average  
31 chromosomal intermarker distance of the covering markers is greater than 10 cM or the equivalent  
32 thereof.

33 10. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a  
34 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
35 the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional

K1 AMENDED SHEET

58

-84

32. An apparatus for localizing a bi-allelic genetic characteristic gene in a species of creatures to a CL-F location as in claim 31, wherein the apparatus comprises a computer, the computer being supplied with proper data and instructions.

33. A process for obtaining genotype data/sample allele frequency data for each bi-allelic marker of a group of two or more bi-allelic covering markers in the chromosomal DNA of one or more individuals of a sample, each individual in the sample being a member of the same species, comprising:

a) determining information on the presence or absence of each allele of each bi-allelic marker of a group of two or more bi-allelic covering markers in the chromosomal DNA of one or more individuals of a sample, a CL-F region being systematically covered by the two or more bi-allelic covering markers, the CL-F region being a collection of points on a two-dimensional plane, the two-dimensional plane having the two orthogonal dimensions of chromosomal location and least common allele frequency; and

b) transforming the information of step a) into genotype data/sample allele frequency data for each marker of the group.

34. A process for obtaining genotype data/sample allele frequency data as in claim 33, wherein the CL-F region is N covered to within a CL-F distance  $\delta$  by the two or more bi-allelic covering markers, so that each point in the region is within the CL-F distance  $\delta$  of N or more of the covering markers, wherein  $\delta$  is equal to about  $[\delta_{CL}, 0.25]$  or the equivalent thereof,  $\delta_{CL}$  is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has been observed between any polymorphisms in any population of the species, N is an integer greater than or equal to 1.

35. A process for obtaining genotype data/sample allele frequency data as in claim 34, wherein the CL-F region includes 81 percent or more of the centerpoints of the matrix centerpoint lattice of a CL-F matrix, the number of cells in the matrix being greater than or equal to three, wherein the matrix has R rows and C columns, each cell of the matrix being of length  $L_{MC}$  and width  $W_{MC}$ , and  $L_{MC}$  being less than or equal to about  $2\delta_{CL}$ , and  $W_{MC}$  being less than or equal to 0.5,  $\delta_{CL}$  is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has been observed between any polymorphisms in any population of the species, there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix, N is an integer greater than or equal to 1; the covering markers being distributed over a chromosomal region of interest, the region of interest being approximately the smallest chromosome interval that contains all of the covering markers, and the covering markers comprising essentially less than all of the polymorphisms in the region of interest.

AMENDED SHEET

PCT/US99/04376

59  
85

1 36. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the  
2 covering markers are substantially nonevenly distributed across a chromosome or a chromosomal  
3 segment.

4 37. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the  
5 covering markers are substantially evenly distributed across a chromosome or a chromosomal segment,  
6 and wherein there is a subgroup of one or more of the covering markers, and each of the markers in the  
7 subgroup is chosen without substantial preference for the least common allele frequency of each of the  
8 markers in the subgroup being close to 0.5, and the number of covering markers in the subgroup is  
9 about 5 percent or more of the total number of covering markers.

10 1038. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the  
11 covering markers are substantially evenly distributed across a chromosome or a chromosomal segment,  
12 and wherein there is a subgroup of one or more of the covering markers, and each of the markers in the  
13 subgroup is chosen without substantial preference for the least common allele frequency of each of the  
14 markers in the subgroup being close to 0.5.

15 39. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the  
16 covering markers are substantially evenly distributed across a chromosome or a chromosomal segment,  
17 wherein (1) the average chromosomal intermarker distance of the covering markers is greater than 2 cM  
18 or the equivalent thereof and the least common allele frequency of one or more of the covering markers  
19 is less than 0.3, or wherein (2) the least common allele frequency of one or more of the covering  
20 markers is less than 0.2, or wherein (3) the average chromosomal intermarker distance of the covering  
21 markers is greater than 10 cM or the equivalent thereof.

22 40. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the  
23 covering markers are substantially evenly distributed across a chromosome or a chromosomal segment,  
24 wherein the average chromosomal intermarker distance of the covering markers is less than or equal to  
25 2 cM or the equivalent thereof, and wherein the conditional probability the covering markers were  
26 chosen essentially randomly from substantially the known set of bi-allelic markers with least common  
27 allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein  
28 the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of  
29 the covering markers, (2) the marker type of each covering marker and (3) there being N or more  
30 covering markers in each cell of 81 percent or more of the cells of the matrix.

31 41. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the  
32 covering markers are substantially evenly distributed across a chromosome or a chromosomal segment,  
33 wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or  
34 the equivalent thereof; and wherein the conditional probability the covering markers were chosen  
35 essentially randomly from substantially the known set of bi-allelic markers with least common allele

AMENDED SHEET

53

74  
100

1 ~~1378~~ One or more copies of a set of oligonucleotides, the set of oligonucleotides being substantially  
2 complementary to a group of two or more bi-allelic covering markers of the same species, wherein the  
3 group of covering markers systematically cover a CL-F region, the CL-F region being a collection of  
4 points on a two-dimensional plane, the two-dimensional plane having the two orthogonal dimensions of  
5 chromosomal location and least common allele frequency.

6 ~~1479~~ One or more copies of a set of oligonucleotides as in claim ~~78~~<sup>13</sup>, wherein the CL-F region is N covered  
7 to within a CL-F distance  $\delta$  by the two or more bi-allelic covering markers, so that each point in the  
8 region is within the CL-F distance  $\delta$  of N or more of the covering markers, wherein  $\delta$  is equal to about [  
9  $\delta_{CL}$ , 0.25] or the equivalent thereof,  $\delta_{CL}$  is equal to the largest chromosomal length, computed by any  
10 method, for which linkage disequilibrium has been observed between any polymorphisms in any  
11 population of the species, N is an integer greater than or equal to 1.

12 ~~1580~~ One or more copies of a set of oligonucleotides as in claim ~~78~~<sup>13</sup>, wherein the CL-F region includes 81  
13 percent or more of the centerpoints of the matrix centerpoint lattice of a CL-F matrix, the number of cells  
14 in the matrix being greater than or equal to three, wherein the matrix has R rows and C columns, each  
15 cell of the matrix being of length  $L_{MC}$  and width  $W_{MC}$ , and  $L_{MC}$  being less than or equal to about  $2\delta_{CL}$ , and  
16  $W_{MC}$  being less than or equal to 0.5,  $\delta_{CL}$  is equal to the largest chromosomal length, computed by any  
17 method, for which linkage disequilibrium has been observed between any polymorphisms in any  
18 population of the species, there being N or more covering markers in each cell of 81 percent or more of  
19 the cells of the matrix, N is an integer greater than or equal to 1; the covering markers being distributed  
20 over a chromosomal region of interest, the region of interest being approximately the smallest  
21 chromosome interval that contains all of the covering markers, and the covering markers comprising  
22 essentially less than all of the polymorphisms in the region of interest.

23 81. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are  
24 substantially nonevenly distributed across a chromosome or a chromosomal segment.

25 82. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are  
26 substantially evenly distributed across a chromosome or a chromosomal segment, and wherein there is  
27 a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen  
28 without substantial preference for the least common allele frequency of each of the markers in the  
29 subgroup being close to 0.5, and the number of covering markers in the subgroup is about 5 percent or  
30 more of the total number of covering markers.

31 ~~1683~~ One or more copies of a set of oligonucleotides as in claim ~~80~~<sup>15</sup>, wherein the covering markers are  
32 substantially evenly distributed across a chromosome or a chromosomal segment, and wherein there is  
33 a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen  
34 without substantial preference for the least common allele frequency of each of the markers in the  
35 subgroup being close to 0.5.

AMENDED SHEET

54

PCT/US99/04376

IPEA/US 22 MAY 2000

51

~~77~~

1 probability the covering markers were chosen essentially randomly from substantially the known set of  
2 bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less  
3 than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the  
4 approximate chromosomal distribution of the covering markers, (2) the marker type of each covering  
5 marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of  
6 the matrix.

7 11. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a  
8 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
9 the covering markers is greater than 2 cM or the equivalent thereof; and wherein the conditional  
10 probability the covering markers were chosen essentially randomly from substantially the known set of  
11 bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less  
12 than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the  
13 approximate chromosomal distribution of the covering markers, (2) the marker type of each covering  
14 marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of  
15 the matrix.

16 12. A process as in claim 10, wherein the chromosome or the chromosomal segment consists  
17 essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein  
18 each chromosome segment of the set has two or less covering markers located thereon, wherein one  
19 and only one covering marker is located on each of 80 percent or more of the chromosome segments of  
20 the set, and wherein zero or two and only two covering markers are located on each of 20 percent or  
21 less of the chromosome segments of the set, and wherein each chromosome segment that borders a  
22 chromosome segment with zero covering markers located thereon has only one or two covering  
23 markers located thereon, and wherein each chromosome segment that borders a chromosome segment  
24 with two covering markers located thereon has only one or zero covering markers located thereon; and  
25 wherein the conditional probability the covering markers were chosen essentially randomly from  
26 substantially the known set of bi-allelic markers with least common allele frequencies between 0.2  
27 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is  
28 substantially conditional on (1) the chromosomal distribution of the covering markers on the  
29 chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or  
30 more covering markers in each cell of 81 percent or more of the cells of the matrix.

31 13. A process as in claim 11, wherein the chromosome or the chromosomal segment consists  
32 essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein  
33 each chromosome segment of the set has two or less covering markers located thereon, wherein one  
34 and only one covering marker is located on each of 80 percent or more of the chromosome segments of  
35 the set, and wherein zero or two and only two covering markers are located on each of 20 percent or

AMENDED SHEET

PCT/US99/04376

IPEA/US 22 MAY 2000

52  
78

less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

14. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells of length  $L_{MC}$  and width  $W_{MC}$  in R rows and C columns; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

15. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele

**AMENDED SHEET**

53

~~79~~

1 frequencies between 0.3 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in  
2 the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal  
3 distribution of the group of markers and the covering markers is substantially similar, and wherein  
4 criterion (4) is, the marker type of each group marker and the covering marker with substantially the  
5 same chromosomal location is the same; wherein a group that is a member of collection C meets  
6 criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more  
7 of the cells of a CL-F matrix with cells of length  $L_{MC}$  and width  $W_{MC}$  in R rows and C columns; wherein P  
8 is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90  
9 percent.

10 16. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a  
11 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
12 the covering markers is less than or equal to 2 cM or the equivalent thereof; wherein the chromosome or  
13 the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of  
14 substantially equal length, and wherein each chromosome segment of the set has two or less covering  
15 markers located thereon, wherein one and only one covering marker is located on each of 80 percent or  
16 more of the chromosome segments of the set, and wherein zero or two and only two covering markers  
17 are located on each of 20 percent or less of the chromosome segments of the set, and wherein each  
18 chromosome segment that borders a chromosome segment with zero covering markers located thereon  
19 has only one or two covering markers located thereon, and wherein each chromosome segment that  
20 borders a chromosome segment with two covering markers located thereon has only one or zero  
21 covering markers located thereon; wherein collection D is essentially the collection of known groups of  
22 bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each  
23 group in the collection being substantially similar to the covering markers as a group; wherein a group of  
24 bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2),  
25 (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set  
26 of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive,  
27 wherein criterion (2) is, the number of markers in the group is the same as the number of covering  
28 markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering  
29 markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and  
30 the covering marker with substantially the same chromosomal location is the same; wherein a group  
31 that is a member of collection D meets criterion (5) if and only if (5) there are N or more of the group  
32 markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells of length  $L_{MC}$  and  
33 width  $W_{MC}$  in R rows and C columns; wherein P is essentially the proportion of groups in collection D  
34 that meet criterion (5); wherein P is less than 90 percent.

**AMENDED SHEET**



IPEA/US 22 MAY 2000

PCT/US99/04376

54  
80

17. A process as in claim 5, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection D meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells having length  $L_{MC}$  and width  $W_{MC}$  in R rows and C columns; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than 90 percent.

18. A process as in claim 4, wherein the covering markers meet criterion a) and one or more of the criteria b), c), d), e), f), g), h) or i), wherein

criterion a) is the covering markers are distributed over a chromosomal region of interest, the region of interest being approximately the smallest chromosome interval that contains all of the covering markers, and the covering markers comprise essentially less than all of the polymorphisms in the region of interest;

criterion b) is the covering markers are substantially nonevenly distributed across a chromosome or a chromosomal segment;

**AMENDED SHEET**

1  
2 criterion c) is the covering markers are substantially evenly distributed across a chromosome or a  
3 chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of  
4 the markers in the subgroup is chosen without substantial preference for the least common allele  
5 frequency of each of the markers in the subgroup being close to 0.5, and the number of covering  
6 markers in the subgroup is about 5 percent or more of the total number of covering markers;

7  
8 criterion d) is the covering markers are substantially evenly distributed across a chromosome or a  
9 chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of  
10 the markers in the subgroup is chosen without substantial preference for the least common allele  
11 frequency of each of the markers in the subgroup being close to 0.5;

12  
13 criterion e) is the covering markers are substantially evenly distributed across a chromosome or a  
14 chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering  
15 markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or  
16 more of the covering markers is less than 0.3, or wherein (2) the least common allele frequency of one  
17 or more of the covering markers is less than 0.2, or wherein (3) the average chromosomal intermarker  
18 distance of the covering markers is greater than 10 cM or the equivalent thereof;

19  
20 criterion f) is the covering markers are substantially evenly distributed across a chromosome or a  
21 chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers  
22 is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional probability the  
23 covering markers were chosen essentially randomly from substantially the known set of bi-allelic  
24 markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or  
25 equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the  
26 approximate chromosomal distribution of the covering markers, (2) the marker type of each covering  
27 marker and (3) the CL-F region being N covered to within the CL-F distance  $\delta$  by the two or more bi-  
28 allelic covering markers;

29  
30 criterion g) is the covering markers are substantially evenly distributed across a chromosome or a  
31 chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers  
32 is greater than 2 cM or the equivalent thereof; and wherein the conditional probability the covering  
33 markers were chosen essentially randomly from substantially the known set of bi-allelic markers with  
34 least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10  
35 percent; wherein the conditional probability is substantially conditional on (1) the approximate

AMENDED SHEET

PCT/US99/04376

PCT/US 99/04376

IPEAUS 22 MAY 2000

56

-82

1 chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3)  
2 the CL-F region being N covered to within the CL-F distance  $\delta$  by the two or more bi-allelic covering  
3 markers;

4  
5 criterion h) is the covering markers are substantially evenly distributed across a chromosome or a  
6 chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers  
7 is less than or equal to 2 cM or the equivalent thereof, and wherein collection C is essentially the  
8 collection of known groups of bi-allelic markers with least common allele frequencies between 0.2  
9 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering  
10 markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the  
11 group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group  
12 is chosen from substantially the known set of bi-allelic markers with least common allele frequencies  
13 between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is  
14 the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of  
15 the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the  
16 marker type of each group marker and the covering marker with substantially the same chromosomal  
17 location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if  
18 (5) the CL-F region is N covered to within a CL-F distance  $\delta$  by the two or more bi-allelic covering  
19 markers; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein  
20 P is less than 90 percent;

21  
22 and criterion i) is the covering markers are substantially evenly distributed across a chromosome or a  
23 chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers  
24 is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of  
25 known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5  
26 inclusive, each group in the collection being substantially similar to the covering markers as a group;  
27 wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially  
28 meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from  
29 substantially the known set of bi-allelic markers with least common allele frequencies between 0.3  
30 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as  
31 the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of  
32 markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of  
33 each group marker and the covering marker with substantially the same chromosomal location is the  
34 same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) the CL-F  
35 region is N covered to within a CL-F distance  $\delta$  by the two or more bi-allelic covering markers; wherein P

AMENDED SHEET

PCT/US 99/04376

PCT/US99/04376

IPEA/US 22 MAY 2000

57

-83

- 1 is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90  
2 percent.
- 3 19. A process as in claim 18, wherein  $\delta$  is less than or equal to about [1 cM, 0.15] or the equivalent  
4 thereof.
- 5 20. A process as in any one of claims 3-19, wherein there is a subgroup of the covering markers, and  
6 the markers in the subgroup are a majority of the covering markers, and each marker in the subgroup is  
7 an SNP, or a bi-allelic marker equivalent formed only from one or more SNPs.
- 8 21. A process as in claim 20, wherein the process comprises a computer program.
- 9 22. A process as in claim 20 wherein  $L_{MC}$  is less than or equal to about 250,000 bp or the equivalent  
10 thereof,  $W_{MC}$  is less than or equal to about 0.15, wherein the species is human being, wherein the same  
11 statistical linkage test based on allelic association is chosen for each covering marker in step b).
- 12 23. A process as in any one of claim 22, wherein the process comprises a computer program.
- 13 24. An apparatus for identifying bi-allelic markers linked to a bi-allelic genetic characteristic gene in a  
14 species of creatures, comprising: means to practice each of the steps of a process as in claim 20.
- 15 25. An apparatus as in claim 24, wherein the apparatus comprises oligonucleotide technology or mass  
16 spectrometry.
- 17 26. An apparatus for identifying bi-allelic markers linked to a bi-allelic genetic characteristic gene in a  
18 species of creatures, comprising: means to practice each of the steps of a process as in any one of the  
19 claims 3-19.
- 20 27. An apparatus as in claim 26, wherein the apparatus comprises a computer, the computer being  
21 supplied with proper data and instructions.
- 22
- 23 28. A process for localizing a bi-allelic genetic characteristic gene in a species of creatures to a CL-F  
24 location, comprising the steps of: a process as in claim 20, further comprising the step f) of: localizing  
25 the gene to the CL-F location of one or more markers that show evidence for linkage based on the  
26 calculations of step e).
- 27 29. A process for localizing a bi-allelic genetic characteristic gene in a species of creatures to a CL-F  
28 location, comprising the steps of: a process in any one of the claims 3-19; further comprising: the step f)  
29 of: localizing the gene to the CL-F location of one or more markers that show evidence for linkage  
30 based on the calculations of step e).
- 31 30. A process as in claim 29, wherein the process comprises a computer program.
- 32 31. An apparatus for localizing a bi-allelic genetic characteristic gene in a species of creatures to a CL-F  
33 location, comprising: means to practice each of the steps of a process as in claim 29.

AMENDED SHEET

40

86-

1 frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the  
2 conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the  
3 covering markers, (2) the marker type of each covering marker and (3) there being N or more covering  
4 markers in each cell of 81 percent or more of the cells of the matrix.

5 42. A process for obtaining genotype data/sample allele frequency data as in claim 40, wherein the  
6 chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome  
7 segments of substantially equal length, and wherein each chromosome segment of the set has two or  
8 less covering markers located thereon, wherein one and only one covering marker is located on each of  
9 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two  
10 covering markers are located on each of 20 percent or less of the chromosome segments of the set,  
11 and wherein each chromosome segment that borders a chromosome segment with zero covering  
12 markers located thereon has only one or two covering markers located thereon, and wherein each  
13 chromosome segment that borders a chromosome segment with two covering markers located thereon  
14 has only one or zero covering markers located thereon; and wherein the conditional probability the  
15 covering markers were chosen essentially randomly from substantially the known set of bi-allelic  
16 markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or  
17 equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the  
18 chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the  
19 marker type of each covering marker and (3) there being N or more covering markers in each cell of 81  
20 percent or more of the cells of the matrix.

21 43. A process for obtaining genotype data/sample allele frequency data as in claim 41, wherein the  
22 chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome  
23 segments of substantially equal length, and wherein each chromosome segment of the set has two or  
24 less covering markers located thereon, wherein one and only one covering marker is located on each of  
25 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two  
26 covering markers are located on each of 20 percent or less of the chromosome segments of the set,  
27 and wherein each chromosome segment that borders a chromosome segment with zero covering  
28 markers located thereon has only one or two covering markers located thereon, and wherein each  
29 chromosome segment that borders a chromosome segment with two covering markers located thereon  
30 has only one or zero covering markers located thereon; and wherein the conditional probability the  
31 covering markers were chosen essentially randomly from substantially the known set of bi-allelic  
32 markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or  
33 equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the  
34 chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the

**AMENDED SHEET**

61  
-87

1 marker type of each covering marker and (3) there being N or more covering markers in each cell of 81  
2 percent or more of the cells of the matrix.

3 44. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the  
4 covering markers are substantially evenly distributed across a chromosome or a chromosomal segment,  
5 wherein the average chromosomal intermarker distance of the covering markers is less than or equal to  
6 2 cM or the equivalent thereof; and wherein collection C is essentially the collection of known groups of  
7 bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each  
8 group in the collection being substantially similar to the covering markers as a group; wherein a group of  
9 bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2),  
10 (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set  
11 of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive,  
12 wherein criterion (2) is, the number of markers in the group is the same as the number of covering  
13 markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering  
14 markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and  
15 the covering marker with substantially the same chromosomal location is the same; wherein a group  
16 that is a member of collection C meets criterion (5) if and only if (5) there are N or more of the group  
17 markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells having length  $L_{MC}$  and  
18 width  $W_{MC}$  in R rows and C columns; wherein P is essentially the proportion of groups in collection C  
19 that meet criterion (5); wherein P is less than 90 percent.

20 45. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the  
21 covering markers are substantially evenly distributed across a chromosome or a chromosomal segment,  
22 wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or  
23 the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic  
24 markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in  
25 the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic  
26 markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and  
27 (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-  
28 allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, wherein  
29 criterion (2) is, the number of markers in the group is the same as the number of covering markers,  
30 wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers  
31 is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the  
32 covering marker with substantially the same chromosomal location is the same; wherein a group that is  
33 a member of collection C meets criterion (5) if and only if (5) there are N or more of the group markers in  
34 each cell of 81 percent or more of the cells of a CL-F matrix with cells having length  $L_{MC}$  and width  $W_{MC}$

**AMENDED SHEET**

62

88

1 in R rows and C columns; wherein P is essentially the proportion of groups in collection C that meet  
2 criterion (5); wherein P is less than 90 percent.

3 46. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the  
4 covering markers are substantially evenly distributed across a chromosome or a chromosomal segment,  
5 wherein the average chromosomal intermarker distance of the covering markers is less than or equal to  
6 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists  
7 essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein  
8 each chromosome segment of the set has two or less covering markers located thereon, wherein one  
9 and only one covering marker is located on each of 80 percent or more of the chromosome segments of  
10 the set, and wherein zero or two and only two covering markers are located on each of 20 percent or  
11 less of the chromosome segments of the set, and wherein each chromosome segment that borders a  
12 chromosome segment with zero covering markers located thereon has only one or two covering  
13 markers located thereon, and wherein each chromosome segment that borders a chromosome segment  
14 with two covering markers located thereon has only one or zero covering markers located thereon;  
15 wherein collection D is essentially the collection of known groups of bi-allelic markers with least common  
16 allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being  
17 substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a  
18 member of collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each  
19 marker in the group is chosen from substantially the known set of bi-allelic markers with least common  
20 allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the number of covering markers and the  
21 number of group markers located on each chromosome segment of the set is the same, and (3) there is  
22 a group marker of the same type as each covering marker located on the same chromosome segment  
23 of the set as each covering marker; wherein a group that is a member of collection D substantially  
24 meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix;  
25 wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less  
26 than 90 percent.

27 47. A process for obtaining genotype data/sample allele frequency data as in claim 35, wherein the  
28 covering markers are substantially evenly distributed across a chromosome or a chromosomal segment,  
29 wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or  
30 the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a  
31 set of nonoverlapping chromosome segments of substantially equal length, and wherein each  
32 chromosome segment of the set has two or less covering markers located thereon, wherein one and  
33 only one covering marker is located on each of 80 percent or more of the chromosome segments of the  
34 set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of  
35 the chromosome segments of the set, and wherein each chromosome segment that borders a

**AMENDED SHEET**

43

~~89~~

1 chromosome segment with zero covering markers located thereon has only one or two covering  
2 markers located thereon, and wherein each chromosome segment that borders a chromosome segment  
3 with two covering markers located thereon has only one or zero covering markers located thereon;  
4 wherein collection D is essentially the collection of known groups of bi-allelic markers with least common  
5 allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being  
6 substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a  
7 member of collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each  
8 marker in the group is chosen from substantially the known set of bi-allelic markers with least common  
9 allele frequencies between 0.3 inclusive and 0.5 inclusive, (2) the number of covering markers and the  
10 number of group markers located on each chromosome segment of the set is the same, and (3) there is  
11 a group marker of the same type as each covering marker located on the same chromosome segment  
12 of the set as each covering marker; wherein a group that is a member of collection D substantially  
13 meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix;  
14 wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less  
15 than 90 percent.

16 48. A process for obtaining genotype data/sample allele frequency data as in claim 34, wherein the  
17 covering markers meet criterion a) and one or more of the criteria b), c), d), e), f), g), h) or i), wherein  
18  
19 criterion a) is the covering markers are distributed over a chromosomal region of interest, the region of  
20 interest being approximately the smallest chromosome interval that contains all of the covering markers,  
21 and the covering markers comprise essentially less than all of the polymorphisms in the region of  
22 interest;

23  
24 criterion b) is the covering markers are substantially nonevenly distributed across a chromosome or a  
25 chromosomal segment;

26  
27 criterion c) is the covering markers are substantially evenly distributed across a chromosome or a  
28 chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of  
29 the markers in the subgroup is chosen without substantial preference for the least common allele  
30 frequency of each of the markers in the subgroup being close to 0.5, and the number of covering  
31 markers in the subgroup is about 5 percent or more of the total number of covering markers;

32  
33 criterion d) is the covering markers are substantially evenly distributed across a chromosome or a  
34 chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of  
35 the markers in the subgroup is chosen without substantial preference for the least common allele

**AMENDED SHEET**



PCT/US 99/04376

PCT/US99/04376

IPEA/US 22 MAY 2000

64  
90

frequency of each of the markers in the subgroup being close to 0.5;

criterion e) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or more of the covering markers is less than 0.3, or wherein (2) the least common allele frequency of one or more of the covering markers is less than 0.2, or wherein (3) the average chromosomal intermarker distance of the covering markers is greater than 10 cM or the equivalent thereof;

criterion f) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) the CL-F region being N covered to within the CL-F distance  $\delta$  by the two or more bi-allelic covering markers;

criterion g) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) the CL-F region being N covered to within the CL-F distance  $\delta$  by the two or more bi-allelic covering markers;

criterion h) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the

**AMENDED SHEET**

PCT/US99/04376

65

91

1 group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group  
2 is chosen from substantially the known set of bi-allelic markers with least common allele frequencies  
3 between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is  
4 the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of  
5 the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the  
6 marker type of each group marker and the covering marker with substantially the same chromosomal  
7 location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if  
8 (5) the CL-F region is N covered to within a CL-F distance  $\delta$  by the two or more bi-allelic covering  
9 markers; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein  
10 P is less than 90 percent;

11  
12 and criterion i) is the covering markers are substantially evenly distributed across a chromosome or a  
13 chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers  
14 is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of  
15 known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5  
16 inclusive, each group in the collection being substantially similar to the covering markers as a group;  
17 wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially  
18 meets criteria (1), (2), (3) and (4), wherein criterion (1) is, each marker in the group is chosen from  
19 substantially the known set of bi-allelic markers with least common allele frequencies between 0.3  
20 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as  
21 the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of  
22 markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of  
23 each group marker and the covering marker with substantially the same chromosomal location is the  
24 same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) the CL-F  
25 region is N covered to within a CL-F distance  $\delta$  by the two or more bi-allelic covering markers; wherein P  
26 is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90  
27 percent.

28 49. A process for obtaining genotype data/sample allele frequency data as in claim 48, wherein  $\delta$  is less  
29 than or equal to about [1 cM, 0.15] or the equivalent thereof.

30 ~~50. A process for obtaining genotype data/sample allele frequency data as in any one of claims 33-49,~~  
31 ~~wherein there is a subgroup of the covering markers, and the markers in the subgroup are a majority of~~  
32 ~~the covering markers, and each marker in the subgroup is an SNP, or a bi-allelic marker equivalent~~  
33 ~~formed only from one or more SNPs.~~

34 51. A process for obtaining genotype data/sample allele frequency data as in claim 50, wherein the  
35 process comprises a computer program.

AMENDED SHEET

66  
92

52. A process for obtaining genotype data/sample allele frequency data as in claim 50 wherein  $L_{MC}$  is less than or equal to about 250,000 bp or the equivalent thereof,  $W_{MC}$  is less than or equal to about 0.15, wherein the species is human being.

53. A process for obtaining genotype data/sample allele frequency data as in claim 52, wherein the process comprises a computer program.

54. An apparatus for obtaining genotype data/sample allele frequency data for each bi-allelic marker of a group of two or more bi-allelic covering markers in the chromosomal DNA of one or more individuals of a sample, each individual in the sample being a member of the same species, comprising: means to practice each of the steps of a process as in any one of the claims 33-49.

55. An apparatus for obtaining genotype data/sample allele frequency data for each bi-allelic marker of a group of two or more bi-allelic covering markers in the chromosomal DNA of one or more individuals of a sample, each individual in the sample being a member of the same species, comprising: means to practice each of the steps of a process as in claim 50.

56. An apparatus as in claim 55, wherein the apparatus comprises oligonucleotide technology or mass spectrometry.

57. An apparatus as in claim 54, wherein the apparatus comprises oligonucleotide technology or mass spectrometry.

58. An apparatus as in claim 54, wherein the apparatus comprises a computer, the computer being supplied with proper data and instructions.

59. A use of one or more copies of a set of oligonucleotides to determine genotype data/sample allele frequency data for each bi-allelic marker of a group of two or more bi-allelic covering markers for one or more individuals, each individual being a member of the same species, wherein the group of covering markers systematically cover a CL-F region, the CL-F region being a collection of points on a two-dimensional plane, the two-dimensional plane having the two orthogonal dimensions of chromosomal location and least common allele frequency.

60. A use as in claim 59, wherein the CL-F region is N covered to within a CL-F distance  $\delta$  by the two or more bi-allelic covering markers, so that each point in the region is within the CL-F distance  $\delta$  of N or more of the covering markers, wherein  $\delta$  is equal to about  $[\delta_{CL}, 0.25]$  or the equivalent thereof,  $\delta_{CL}$  is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium has been observed between any polymorphisms in any population of the species, N is an integer greater than or equal to 1.

61. A use as in claim 59, wherein the CL-F region includes 81 percent or more of the centerpoints of the matrix centerpoint lattice of a CL-F matrix, the number of cells in the matrix being greater than or equal

**AMENDED SHEET**

US 99/04376

IPEAUS 22 MAY 2000

PCT/US99/04376

67  
93

1 to three, wherein the matrix has  $R$  rows and  $C$  columns, each cell of the matrix being of length  $L_{MC}$  and  
2 width  $W_{MC}$ , and  $L_{MC}$  being less than or equal to about  $2\delta_{CL}$ , and  $W_{MC}$  being less than or equal to  $0.5 \cdot \delta_{CL}$   
3 is equal to the largest chromosomal length, computed by any method, for which linkage disequilibrium  
4 has been observed between any polymorphisms in any population of the species, there being  $N$  or more  
5 covering markers in each cell of 81 percent or more of the cells of the matrix,  $N$  is an integer greater  
6 than or equal to 1; the covering markers being distributed over a chromosomal region of interest, the  
7 region of interest being approximately the smallest chromosome interval that contains all of the covering  
8 markers, and the covering markers comprising essentially less than all of the polymorphisms in the  
9 region of interest.

10 62. A use as in claim 61, wherein the covering markers are substantially nonevenly distributed across a  
11 chromosome or a chromosomal segment.

12 63. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a  
13 chromosome or a chromosomal segment, and wherein there is a subgroup of one or more of the  
14 covering markers, and each of the markers in the subgroup is chosen without substantial preference for  
15 the least common allele frequency of each of the markers in the subgroup being close to 0.5, and the  
16 number of covering markers in the subgroup is about 5 percent or more of the total number of covering  
17 markers.

18 64. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a  
19 chromosome or a chromosomal segment, and wherein there is a subgroup of one or more of the  
20 covering markers, and each of the markers in the subgroup is chosen without substantial preference for  
21 the least common allele frequency of each of the markers in the subgroup being close to 0.5.

22 65. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a  
23 chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance of  
24 the covering markers is greater than 2 cM or the equivalent thereof and the least common allele  
25 frequency of one or more of the covering markers is less than 0.3, or wherein (2) the least common  
26 allele frequency of one or more of the covering markers is less than 0.2, or wherein (3) the average  
27 chromosomal intermarker distance of the covering markers is greater than 10 cM or the equivalent  
28 thereof.

29 66. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a  
30 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
31 the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional  
32 probability the covering markers were chosen essentially randomly from substantially the known set of  
33 bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less  
34 than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the  
35 approximate chromosomal distribution of the covering markers, (2) the marker type of each covering

**AMENDED SHEET**

T/US 99/04376  
IPEA/US 22 MAY 2000

PCT/US99/04376

48

94

1 marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of  
2 the matrix.

3 67. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a  
4 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
5 the covering markers is greater than 2 cM or the equivalent thereof; and wherein the conditional  
6 probability the covering markers were chosen essentially randomly from substantially the known set of  
7 bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less  
8 than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the  
9 approximate chromosomal distribution of the covering markers, (2) the marker type of each covering  
10 marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of  
11 the matrix.

12 68. A use as in claim 66, wherein the chromosome or the chromosomal segment consists essentially of  
13 a set of nonoverlapping chromosome segments of substantially equal length, and wherein each  
14 chromosome segment of the set has two or less covering markers located thereon, wherein one and  
15 only one covering marker is located on each of 80 percent or more of the chromosome segments of the  
16 set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of  
17 the chromosome segments of the set, and wherein each chromosome segment that borders a  
18 chromosome segment with zero covering markers located thereon has only one or two covering  
19 markers located thereon, and wherein each chromosome segment that borders a chromosome segment  
20 with two covering markers located thereon has only one or zero covering markers located thereon; and  
21 wherein the conditional probability the covering markers were chosen essentially randomly from  
22 substantially the known set of bi-allelic markers with least common allele frequencies between 0.2  
23 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is  
24 substantially conditional on (1) the chromosomal distribution of the covering markers on the  
25 chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or  
26 more covering markers in each cell of 81 percent or more of the cells of the matrix.

27 69. A use as in claim 67, wherein the chromosome or the chromosomal segment consists essentially of  
28 a set of nonoverlapping chromosome segments of substantially equal length, and wherein each  
29 chromosome segment of the set has two or less covering markers located thereon, wherein one and  
30 only one covering marker is located on each of 80 percent or more of the chromosome segments of the  
31 set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of  
32 the chromosome segments of the set, and wherein each chromosome segment that borders a  
33 chromosome segment with zero covering markers located thereon has only one or two covering  
34 markers located thereon, and wherein each chromosome segment that borders a chromosome segment  
35 with two covering markers located thereon has only one or zero covering markers located thereon; and

**AMENDED SHEET**

69  
85

wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

70. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof; and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar; and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the same chromosomal location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more of the cells of a CL-F matrix with cells having length  $L_{MC}$  and width  $W_{MC}$  in R rows and C columns; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90 percent.

71. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with substantially the

**AMENDED SHEET**

70  
96

1 same chromosomal location is the same; wherein a group that is a member of collection C meets  
2 criterion (5) if and only if (5) there are N or more of the group markers in each cell of 81 percent or more  
3 of the cells of a CL-F matrix with cells having length  $L_{MC}$  and width  $W_{MC}$  in R rows and C columns;  
4 wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less  
5 than 90 percent.

6 72. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a  
7 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
8 the covering markers is less than or equal to 2 cM or the equivalent thereof; wherein the chromosome or  
9 the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of  
10 substantially equal length, and wherein each chromosome segment of the set has two or less covering  
11 markers located thereon, wherein one and only one covering marker is located on each of 80 percent or  
12 more of the chromosome segments of the set, and wherein zero or two and only two covering markers  
13 are located on each of 20 percent or less of the chromosome segments of the set, and wherein each  
14 chromosome segment that borders a chromosome segment with zero covering markers located thereon  
15 has only one or two covering markers located thereon, and wherein each chromosome segment that  
16 borders a chromosome segment with two covering markers located thereon has only one or zero  
17 covering markers located thereon; wherein collection D is essentially the collection of known groups of  
18 bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each  
19 group in the collection being substantially similar to the covering markers as a group; wherein a group of  
20 bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2),  
21 and (3): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers  
22 with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the number of  
23 covering markers and the number of group markers located on each chromosome segment of the set is  
24 the same, and (3) there is a group marker of the same type as each covering marker located on the  
25 same chromosome segment of the set as each covering marker; wherein a group that is a member of  
26 collection D substantially meets criterion (5) if and only if (5) there are N or more of the group markers in  
27 each cell of the matrix; wherein P is essentially the proportion of groups in collection D that meet  
28 criterion (5); wherein P is less than 90 percent.

29 73. A use as in claim 61, wherein the covering markers are substantially evenly distributed across a  
30 chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of  
31 the covering markers is greater than 2 cM or the equivalent thereof; wherein the chromosome or the  
32 chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of  
33 substantially equal length, and wherein each chromosome segment of the set has two or less covering  
34 markers located thereon, wherein one and only one covering marker is located on each of 80 percent or  
35 more of the chromosome segments of the set, and wherein zero or two and only two covering markers

**AMENDED SHEET**

71  
97

are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, (2) the number of covering markers and the number of group markers located on each chromosome segment of the set is the same, and (3) there is a group marker of the same type as each covering marker located on the same chromosome segment of the set as each covering marker; wherein a group that is a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than 90 percent.

74. A use as in claim 60, wherein the covering markers meet criterion a) and one or more of the criteria b), c), d), e), f), g), h) or i), wherein

criterion a) is the covering markers are distributed over a chromosomal region of interest, the region of interest being approximately the smallest chromosome interval that contains all of the covering markers, and the covering markers comprise essentially less than all of the polymorphisms in the region of interest;

criterion b) is the covering markers are substantially nonevenly distributed across a chromosome or a chromosomal segment;

criterion c) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of the markers in the subgroup is chosen without substantial preference for the least common allele frequency of each of the markers in the subgroup being close to 0.5, and the number of covering markers in the subgroup is about 5 percent or more of the total number of covering markers;

criterion d) is the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of

**AMENDED SHEET**



72  
98

1 the markers in the subgroup is chosen without substantial preference for the least common allele  
2 frequency of each of the markers in the subgroup being close to 0.5;

3  
4 criterion e) is the covering markers are substantially evenly distributed across a chromosome or a  
5 chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering  
6 markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or  
7 more of the covering markers is less than 0.3, or wherein (2) the least common allele frequency of one  
8 or more of the covering markers is less than 0.2, or wherein (3) the average chromosomal intermarker  
9 distance of the covering markers is greater than 10 cM or the equivalent thereof;

10  
11 criterion f) is the covering markers are substantially evenly distributed across a chromosome or a  
12 chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers  
13 is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional probability the  
14 covering markers were chosen essentially randomly from substantially the known set of bi-allelic  
15 markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or  
16 equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the  
17 approximate chromosomal distribution of the covering markers, (2) the marker type of each covering  
18 marker and (3) the CL-F region being N covered to within the CL-F distance  $\delta$  by the two or more bi-  
19 allelic covering markers;

20  
21 criterion g) is the covering markers are substantially evenly distributed across a chromosome or a  
22 chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers  
23 is greater than 2 cM or the equivalent thereof; and wherein the conditional probability the covering  
24 markers were chosen essentially randomly from substantially the known set of bi-allelic markers with  
25 least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10  
26 percent; wherein the conditional probability is substantially conditional on (1) the approximate  
27 chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3)  
28 the CL-F region being N covered to within the CL-F distance  $\delta$  by the two or more bi-allelic covering  
29 markers;

30  
31 criterion h) is the covering markers are substantially evenly distributed across a chromosome or a  
32 chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers  
33 is less than or equal to 2 cM or the equivalent thereof; and wherein collection C is essentially the  
34 collection of known groups of bi-allelic markers with least common allele frequencies between 0.2  
35 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering

**AMENDED SHEET**

73

89

1 markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the  
2 group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group  
3 is chosen from substantially the known set of bi-allelic markers with least common allele frequencies  
4 between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is  
5 the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of  
6 the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the  
7 marker type of each group marker and the covering marker with substantially the same chromosomal  
8 location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if  
9 (5) the CL-F region is N covered to within a CL-F distance  $\delta$  by the two or more bi-allelic covering  
10 markers; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein  
11 P is less than 90 percent;

12  
13 and criterion i) is the covering markers are substantially evenly distributed across a chromosome or a  
14 chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers  
15 is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of  
16 known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5  
17 inclusive, each group in the collection being substantially similar to the covering markers as a group;  
18 wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially  
19 meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from  
20 substantially the known set of bi-allelic markers with least common allele frequencies between 0.3  
21 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as  
22 the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of  
23 markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of  
24 each group marker and the covering marker with substantially the same chromosomal location is the  
25 same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) the CL-F  
26 region is N covered to within a CL-F distance  $\delta$  by the two or more bi-allelic covering markers; wherein P  
27 is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90  
28 percent.

29 75. A use as in claim 74, wherein  $\delta$  is less than or equal to about [1 cM, 0.15] or the equivalent thereof.

30 76. A use as in any one of claims 59-75, wherein there is a subgroup of the covering markers, and the  
31 markers in the subgroup are a majority of the covering markers, and each marker in the subgroup is an  
32 SNP, or a bi-allelic marker equivalent formed only from one or more SNPs.

33 77. A use as in claim 76 wherein  $L_{MC}$  is less than or equal to about 250,000 bp or the equivalent thereof,  
34  $W_{MC}$  is less than or equal to about 0.15, wherein the species is human being.

35  
**AMENDED SHEET**

75

101

84. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or more of the covering markers is less than 0.3, or wherein (2) the least common allele frequency of one or more of the covering markers is less than 0.2, or wherein (3) the average chromosomal intermarker distance of the covering markers is greater than 10 cM or the equivalent thereof.

85. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

86. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; and wherein the conditional probability the covering markers were chosen essentially randomly from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the approximate chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3) there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.

87. One or more copies of a set of oligonucleotides as in claim 85, wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; and wherein the conditional probability the covering markers were

**AMENDED SHEET**

76  
102

1 chosen essentially randomly from substantially the known set of bi-allelic markers with least common  
2 allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein  
3 the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering  
4 markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3)  
5 there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.  
6 88. One or more copies of a set of oligonucleotides as in claim 86, wherein the chromosome or the  
7 chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of  
8 substantially equal length, and wherein each chromosome segment of the set has two or less covering  
9 markers located thereon, wherein one and only one covering marker is located on each of 80 percent or  
10 more of the chromosome segments of the set, and wherein zero or two and only two covering markers  
11 are located on each of 20 percent or less of the chromosome segments of the set, and wherein each  
12 chromosome segment that borders a chromosome segment with zero covering markers located thereon  
13 has only one or two covering markers located thereon; and wherein each chromosome segment that  
14 borders a chromosome segment with two covering markers located thereon has only one or zero  
15 covering markers located thereon; and wherein the conditional probability the covering markers were  
16 chosen essentially randomly from substantially the known set of bi-allelic markers with least common  
17 allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10 percent; wherein  
18 the conditional probability is substantially conditional on (1) the chromosomal distribution of the covering  
19 markers on the chromosome segments of the set, (2) the marker type of each covering marker and (3)  
20 there being N or more covering markers in each cell of 81 percent or more of the cells of the matrix.  
21 89. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are  
22 substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average  
23 chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the  
24 equivalent thereof; and wherein collection C is essentially the collection of known groups of bi-allelic  
25 markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, each group in  
26 the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic  
27 markers is a member of collection C if and only if the group substantially meets criteria (1), (2), (3) and  
28 (4); wherein criterion (1) is, each marker in the group is chosen from substantially the known set of bi-  
29 allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, wherein  
30 criterion (2) is, the number of markers in the group is the same as the number of covering markers,  
31 wherein criterion (3) is, the chromosomal distribution of the group of markers and the covering markers  
32 is substantially similar, and wherein criterion (4) is, the marker type of each group marker and the  
33 covering marker with substantially the same chromosomal location is the same; wherein a group that is  
34 a member of collection C meets criterion (5) if and only if (5) there are N or more of the group markers in  
35 each cell of 81 percent or more of the cells of a CL-F matrix with cells having length  $L_{MC}$  and width  $W_{MC}$

**AMENDED SHEET**

77  
103

1 in R rows and C columns; wherein P is essentially the proportion of groups in collection C that meet  
2 criterion (5); wherein P is less than 90 percent.

3 90. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are  
4 substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average  
5 chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent  
6 thereof, and wherein collection C is essentially the collection of known groups of bi-allelic markers with  
7 least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection  
8 being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a  
9 member of collection C if and only if the group substantially meets criteria (1), (2), (3) and (4); wherein  
10 criterion (1) is, each marker in the group is chosen from substantially the known set of bi-allelic markers  
11 with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, wherein criterion (2) is,  
12 the number of markers in the group is the same as the number of covering markers, wherein criterion  
13 (3) is, the chromosomal distribution of the group of markers and the covering markers is substantially  
14 similar, and wherein criterion (4) is, the marker type of each group marker and the covering marker with  
15 substantially the same chromosomal location is the same; wherein a group that is a member of  
16 collection C meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of  
17 81 percent or more of the cells of a CL-F matrix with cells having length  $L_{MC}$  and width  $W_{MC}$  in R rows  
18 and C columns; wherein P is essentially the proportion of groups in collection C that meet criterion (5);  
19 wherein P is less than 90 percent.

20 91. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are  
21 substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average  
22 chromosomal intermarker distance of the covering markers is less than or equal to 2 cM or the  
23 equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set  
24 of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome  
25 segment of the set has two or less covering markers located thereon, wherein one and only one  
26 covering marker is located on each of 80 percent or more of the chromosome segments of the set, and  
27 wherein zero or two and only two covering markers are located on each of 20 percent or less of the  
28 chromosome segments of the set, and wherein each chromosome segment that borders a chromosome  
29 segment with zero covering markers located thereon has only one or two covering markers located  
30 thereon, and wherein each chromosome segment that borders a chromosome segment with two  
31 covering markers located thereon has only one or zero covering markers located thereon; wherein  
32 collection D is essentially the collection of known groups of bi-allelic markers with least common allele  
33 frequencies between 0.2 inclusive and 0.5 inclusive, each group in the collection being substantially  
34 similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of  
35 collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each marker in the

IPEA/US 22 MAY 2000

78

104

group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive, (2) the number of covering markers and the number of group markers located on each chromosome segment of the set is the same, and (3) there is a group marker of the same type as each covering marker located on the same chromosome segment of the set as each covering marker; wherein a group that is a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than 90 percent.

92. One or more copies of a set of oligonucleotides as in claim 80, wherein the covering markers are substantially evenly distributed across a chromosome or a chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers is greater than 2 cM or the equivalent thereof; wherein the chromosome or the chromosomal segment consists essentially of a set of nonoverlapping chromosome segments of substantially equal length, and wherein each chromosome segment of the set has two or less covering markers located thereon, wherein one and only one covering marker is located on each of 80 percent or more of the chromosome segments of the set, and wherein zero or two and only two covering markers are located on each of 20 percent or less of the chromosome segments of the set, and wherein each chromosome segment that borders a chromosome segment with zero covering markers located thereon has only one or two covering markers located thereon, and wherein each chromosome segment that borders a chromosome segment with two covering markers located thereon has only one or zero covering markers located thereon; wherein collection D is essentially the collection of known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering markers as a group; wherein a group of bi-allelic markers is a member of collection D if and only if the group substantially meets criteria (1), (2), and (3): (1) each marker in the group is chosen from substantially the known set of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5 inclusive, (2) the number of covering markers and the number of group markers located on each chromosome segment of the set is the same, and (3) there is a group marker of the same type as each covering marker located on the same chromosome segment of the set as each covering marker; wherein a group that is a member of collection D substantially meets criterion (5) if and only if (5) there are N or more of the group markers in each cell of the matrix; wherein P is essentially the proportion of groups in collection D that meet criterion (5); wherein P is less than 90 percent.

93. One or more copies of a set of oligonucleotides as in claim 79, wherein the covering markers meet criterion a) and one or more of the criteria b), c), d), e), f), g), h) or i), wherein

**AMENDED SHEET**

79

105

1 criterion a) is the covering markers are distributed over a chromosomal region of interest, the region of  
2 interest being approximately the smallest chromosome interval that contains all of the covering markers,  
3 and the covering markers comprise essentially less than all of the polymorphisms in the region of  
4 interest;

6 criterion b) is the covering markers are substantially nonevenly distributed across a chromosome or a  
7 chromosomal segment;

9 criterion c) is the covering markers are substantially evenly distributed across a chromosome or a  
10 chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of  
11 the markers in the subgroup is chosen without substantial preference for the least common allele  
12 frequency of each of the markers in the subgroup being close to 0.5, and the number of covering  
13 markers in the subgroup is about 5 percent or more of the total number of covering markers;

15 criterion d) is the covering markers are substantially evenly distributed across a chromosome or a  
16 chromosomal segment, and there is a subgroup of one or more of the covering markers, and each of  
17 the markers in the subgroup is chosen without substantial preference for the least common allele  
18 frequency of each of the markers in the subgroup being close to 0.5;

20 criterion e) is the covering markers are substantially evenly distributed across a chromosome or a  
21 chromosomal segment, wherein (1) the average chromosomal intermarker distance of the covering  
22 markers is greater than 2 cM or the equivalent thereof and the least common allele frequency of one or  
23 more of the covering markers is less than 0.3, or wherein (2) the least common allele frequency of one  
24 or more of the covering markers is less than 0.2, or wherein (3) the average chromosomal intermarker  
25 distance of the covering markers is greater than 10 cM or the equivalent thereof;

27 criterion f) is the covering markers are substantially evenly distributed across a chromosome or a  
28 chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers  
29 is less than or equal to 2 cM or the equivalent thereof, and wherein the conditional probability the  
30 covering markers were chosen essentially randomly from substantially the known set of bi-allelic  
31 markers with least common allele frequencies between 0.2 inclusive and 0.5 inclusive is less than or  
32 equal to 10 percent; wherein the conditional probability is substantially conditional on (1) the  
33 approximate chromosomal distribution of the covering markers, (2) the marker type of each covering  
34 marker and (3) the CL-F region being N covered to within the CL-F distance  $\delta$  by the two or more bi-  
35 allelic covering markers;

**AMENDED SHEET**

80  
406

1  
2 criterion g) is the covering markers are substantially evenly distributed across a chromosome or a  
3 chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers  
4 is greater than 2 cM or the equivalent thereof; and wherein the conditional probability the covering  
5 markers were chosen essentially randomly from substantially the known set of bi-allelic markers with  
6 least common allele frequencies between 0.3 inclusive and 0.5 inclusive is less than or equal to 10  
7 percent; wherein the conditional probability is substantially conditional on (1) the approximate  
8 chromosomal distribution of the covering markers, (2) the marker type of each covering marker and (3)  
9 the CL-F region being N covered to within the CL-F distance  $\delta$  by the two or more bi-allelic covering  
10 markers;

11  
12 criterion h) is the covering markers are substantially evenly distributed across a chromosome or a  
13 chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers  
14 is less than or equal to 2 cM or the equivalent thereof; and wherein collection C is essentially the  
15 collection of known groups of bi-allelic markers with least common allele frequencies between 0.2  
16 inclusive and 0.5 inclusive, each group in the collection being substantially similar to the covering  
17 markers as a group; wherein a group of bi-allelic markers is a member of collection C if and only if the  
18 group substantially meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group  
19 is chosen from substantially the known set of bi-allelic markers with least common allele frequencies  
20 between 0.2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is  
21 the same as the number of covering markers, wherein criterion (3) is, the chromosomal distribution of  
22 the group of markers and the covering markers is substantially similar, and wherein criterion (4) is, the  
23 marker type of each group marker and the covering marker with substantially the same chromosomal  
24 location is the same; wherein a group that is a member of collection C meets criterion (5) if and only if  
25 (5) the CL-F region is N covered to within a CL-F distance  $\delta$  by the two or more bi-allelic covering  
26 markers; wherein P is essentially the proportion of groups in collection C that meet criterion (5); wherein  
27 P is less than 90 percent;

28  
29 and criterion i) is the covering markers are substantially evenly distributed across a chromosome or a  
30 chromosomal segment, wherein the average chromosomal intermarker distance of the covering markers  
31 is greater than 2 cM or the equivalent thereof, and wherein collection C is essentially the collection of  
32 known groups of bi-allelic markers with least common allele frequencies between 0.3 inclusive and 0.5  
33 inclusive, each group in the collection being substantially similar to the covering markers as a group;  
34 wherein a group of bi-allelic markers is a member of collection C if and only if the group substantially  
35 meets criteria (1), (2), (3) and (4); wherein criterion (1) is, each marker in the group is chosen from

**AMENDED SHEET**



81

107

1 substantially the known set of bi-allelic markers with least common allele frequencies between 0.3  
2 inclusive and 0.5 inclusive, wherein criterion (2) is, the number of markers in the group is the same as  
3 the number of covering markers, wherein criterion (3) is, the chromosomal distribution of the group of  
4 markers and the covering markers is substantially similar, and wherein criterion (4) is, the marker type of  
5 each group marker and the covering marker with substantially the same chromosomal location is the  
6 same; wherein a group that is a member of collection C meets criterion (5) if and only if (5) the CL-F  
7 region is N covered to within a CL-F distance  $\delta$  by the two or more bi-allelic covering markers; wherein P  
8 is essentially the proportion of groups in collection C that meet criterion (5); wherein P is less than 90  
9 percent.

10 94. One or more copies of a set of oligonucleotides as in claim 93, wherein  $\delta$  is less than or equal to  
11 about [1 cM, 0.15] or the equivalent thereof.

12 95. One or more copies of a set of oligonucleotides as in any one of claims 78-94, wherein there is a  
13 subgroup of the covering markers, and the markers in the subgroup are a majority of the covering  
14 markers, and each marker in the subgroup is an SNP, or a bi-allelic marker equivalent formed only from  
15 one or more SNPs.

16 96. One or more copies of a set of oligonucleotides as in claim 95 wherein  $L_{MC}$  is less than or equal to  
17 about 250,000 bp or the equivalent thereof,  $W_{MC}$  is less than or equal to about 0.15, wherein the species  
18 is human being.

19  
20  
21  
22

add BF  
add BF

ADD  
C67

AMENDED SHEET